Our research complements previous studies exploring interindividual variability in visuospatial transformations by examining the early stages of perceptual and attentional processes between men and women during mental rotation. Using ERP topographic analyses, we provide important insights about the electrical dynamics underpinning interindividual differences in demonstrating a significant link between early attentional processes and performance scores.

Interindividual differences in brain dynamics of early visual processes: impact on score accuracy in mental rotation task

Paolo Ruggeri[a], Nam Nguyen[a], Alan J. Pegna[b] and Catherine Brandner[a]*

[a] Brain Electrophysiology Attention Movement Laboratory, Institute of Psychology, University of Lausanne, Switzerland.

[b] School of Psychology, The University of Queensland, Brisbane, Qld-4072, Australia.

* Corresponding author: Catherine.Brandner@unil.ch

ABSTRACT

Interindividual variations in the ability to perform visuospatial mental transformations have been investigated extensively, in particular through mental rotation tasks. However, the impact of early visual processes on performance has been largely ignored. To clarify this issue, we explored the time-course of early visual processing (from 0 to 450 ms post-stimulus) using ERP topographic analyses. The main findings demonstrated a significant link between early attentional processes and accuracy scores occurring more than five seconds later, as well as a strong association between spatial covariance and microstate topographies exhibiting substantial genders differences. More specifically, the results indicated that, in a classical mental rotation task, the male brain expends more time processing visual spatial information resulting in a longer bilateral positive potential at posterior-occipital sites. In comparison, the female brain initiates an earlier processing of non-spatial information resulting in a faster transition from a bilateral positive potential of posterior-occipital sites to a negative potential at central-frontal sites. These findings illustrate how a more complete utilization of the spatiotemporal information contained in EEG recordings can provide important insights about the impact of early visual processes on interindividual differences, particularly across gender, and thus shed new light on alternate cognitive strategies.

1. INTRODUCTION

The ability to mentally change the orientation of an object in order to judge its similarity with another object is probably one of the most remarkable demonstrations of visuospatial mental transformation. The exploration of how people solve such problems, operationalized using mental rotation tasks, has highlighted mainly variations in performance between individuals. Interindividual variations arise from many different sources including practice, stimulus attributes or the nature of the comparison task (Pylyshyn, 1979 for a review). Different factors such as the level of anxiety, self-reported confidence or guessing behavior have also been suggested to play a part (e.g., Blough & Slavin, 1987; Goldstein, Haldane, & Mitchell, 1990; Delgado & Prieto, 1996; Cooke-Simpson & Voyer, 2007). In addition to these factors, mental re-orientation of an object's viewpoint can be based on the movement of the object around its axis (mental rotation) or on a mental change in the observer's position with respect to the object (change of perspective) (e.g. Zacks et al, 2000). More relevant to the context of this study, it has been shown that gender – regarded here as an archetypal example of interindividual differences – has a significant impact on the performance of mental rotation (Voyer, Voyer, & Bryden, 1995 for a review). Gender differences in mental rotation are often interpreted in terms of differences in cognitive strategies used to process information (Just & Carpenter, 1976 for a review). It has been suggested that women tend to use a more sequential approach by separately rotating parts of the object, while men tend to use a more holistic approach rotating whole objects (for a review see, Pletzer, 2014). Conforming to this idea, sequential processing should take longer than the holistic one. Though appealing, this suggestion has not yet been firmly confirmed by investigations. For instance, the speed-accuracy tradeoff, investigated by varying the time given for the response, remains equivocal (e.g. Debelak, Gittler, & Arendasy, 2014) as a meta-analysis revealed that differences across gender were significantly larger when the task was administered with time constraints, compared to situations where no such constraints

were present (Voyer, 2011).

Although these reports put forward interesting questions, they do not explain the underlying cognitive processes that might give rise to differences across gender. The effectiveness of mentally transforming visuospatial information can be seen as the ability to perceive and visualize the characteristics of objects. As the retinal image is swept away by a continuous flow of visual stimulation, attentional processes could function as a tuner, filtering information selectively for further processing. As a consequence, attention can be viewed as a function biasing information processing towards what is relevant, and thus affect efficiency in visuospatial mental imagery. With its high temporal precision, electroencephalographic recording provides a tool for clarifying how attention can modulate visuospatial abilities. Visual attention has been extensively studied using visual search tasks and spatial cueing paradigms. Seminal studies of attentional processing highlighted the P1/N1 event-related potentials (ERP) components as those modulated by visual spatial attention (Luck, et al., 1990; Mangun, & Hillyard,1991; Luck, Fan, & Hillyard, 1993; Luck, & Hillyard, 1994; Mangun, 1995; Luck, & Ford, 1998; for reviews, Luck, & Girelli, 1999; Luck, Woodman, & Vogel, 2000). They characterized voluntary attention as an early spatial filter capable of altering higher stages of visual analysis involved in object perception and recognition. Subsequent studies of visual attention processes and P1/N1 components have extended their involvement to the processing of non-spatial visual tasks including basic features, shapes, and faces (e.g. Taylor, 2002; Itier, & Taylor, 2004; Zani, & Proverbio, 2012; Zani, et al, 2015). Although gender differences in electrophysiological measurements were explored some time ago (Shearer, et al., 1984), only a few investigations examined the impact of early visual attention processes on this procedure. In a task requiring participants to make a standard/deviant distinction between stimuli, no differences were observed across gender in the N1, but women elicited larger P2 amplitudes and shorter latencies than men, suggesting that women were faster in detecting the occurrence

of deviant stimuli and directed more attentional resources to these deviant features compared to men (Yuan, et al., 2008). The study of attention shifting in a spatial cueing task showed larger P1 and N1 amplitudes in women compared to men (Feng, et al., 2011). Other studies including mental rotation tasks reported alternatively the presence (Jaušovec, 2012) and the absence of gender differences in the latencies and amplitudes of the P1/N1 components reflecting attentional process (Jaušovec, & Jaušovec, 2009; Beste, Heil, & Konrad, 2010). Indeed, Jaušovec (2012) carried out a mental rotation task in a group of male and female participants matched for intelligence and spatial ability, and observed shorter P1 latencies and more negative N1 components for females in the task. By contrast, Beste et al., found no difference across gender in a mental rotation task involving alphanumeric characters, despite their large group of participants. Taken together, these studies have yielded mixed results ranging from the absence, to the presence of interindividual differences in P1/N1 components linked to visual perception and attentional processes.

Although these findings are invaluable, they suffer from the recourse to selected electrode sites, providing ERP waveforms that depend on the choice of active and reference electrodes used. Because ERP waveforms at a given electrode can be composed of signals produced by sources at various locations, observations at the scalp level can result from either a change of source strength or distribution (Lehmann, & Skrandies, 1980). Thus, interindividual comparisons at predefined time points or temporal windows, and at selected electrode locations, provided valid but incomplete evidence about how early attention processes may impact on accuracy in tasks involving visuospatial mental transformation.

The present work is an attempt to bridge this gap by using a more detailed approach than the traditional ERP amplitudes and latency analysis usually employed when assessing visual perception and attention. More specifically, we aimed to examine differences in the time course of electrical activation between men and women during mental rotation, with a particular

emphasis on the early stages of perceptual and attentional processing corresponding to the classical P1/N1 components. We hypothesized that differences across gender might be linked to early attentional processes which would be reflected on the duration of the transient brain states elicited during visual spatial information processing. To investigate this hypothesis, women and men – regarded here as a typical example of interindividual differences – were subjected to a classical mental rotation task (MR) using Shepard and Metzler stimuli (1971) and to a perspective change task (PT) based on Piaget's 'Three-Mountains Task' (Piaget & Inhelder, 1956). As we did not find similar studies, the PT task was only explored as an alternative mode of mental rotation. To avoid a possible gender bias favoring men, the tasks were administered without time-limited responses. To explore the brain dynamics, ERPs were characterized by assessing the temporal sequence of electrical potential field maps using a combination of complementary methods (TANOVA and microstate analysis) sensitive to the temporality of events. To ensure the association between early attentional processes and response accuracy, we used topographic analysis of covariance (TANCOVA) on the whole sample of data of each task (mental rotation and perspective taking) respectively (Koenig, et al, 2008). To our knowledge, this is the first time that the spatiotemporal information available in EEG recordings is fully exploited to explore the early stages of visual processing in mental rotation tasks.

2. METHOD

2.1 *Participants*

Sixty-seven healthy participants (32 women; mean age 22.9 years) from the University of Lausanne provided informed consent and participated in exchange for course credit. All participants were right-handed, reported no history of neurological diseases and had normal or corrected-to-normal vision. One significant outlier (a man who scored below 25% correct

responses) was eliminated from the data set. Furthermore, 12 subjects (6 women, 6 men) were excluded as it was impossible to compute their ERPs. Indeed, these participants repetitively produced high amplitude saccadic movements and blinks over the entire duration of the task corrupting the EEG epochs with artefacts. The final sample was composed of 29 men and 26 women. Research was approved by the Cantonal Ethics Committee for Human Research (Vaud, Switzerland; protocol N°286/13) and was carried out in accordance with the code of ethics of the World Medical Association (Declaration of Helsinki) for experiments involving human subjects in research.

2.2 *Visual stimuli and task*

Both mental rotation (MR) and perspective taking (PT) 3-D stimuli were drawn using Google SketchUp software and displayed with E-Prime Software (Version2.0, Psychology Software Tools). In the MR task, the 3D objects (FIG. 1A) were similar to those used by Shepard and Metzler (1971). The stimuli of the PT task (FIG. 1B) were drawings of the Piaget's 'Three-Mountain Task' (Piaget & Inhelder, 1956). Stimuli were presented in pairs in both tasks. The stimuli's size and a central fixation cross, equidistant from left and right images, were chosen to decrease eye movements (e.g., saccades). Half of the MR and PT paired objects (right) were identical to the sample (left). In the MR task the remaining stimuli were mirror images while, in the PT task, the remaining half were images of non-matching views. To perform the MR task, participants were instructed to rotate the object on the left around its axis until it matched the stimulus on the right. To perform the PT task, they had to imagine how the three objects in the image on the left would be viewed by an observer situated at a given position (symbolized by the eye on the image) and then had to judge whether the image on the right coincided with the view from the observer's location.

INSERT FIGURE 1 ABOUT HERE

2.3 *Experimental Procedure*

Participants were comfortably seated in a moderately dark room (Faraday cage) with their head placed on a chin rest and facing a monitor placed at 60 cm from their eyes. To minimize eye movements, they were informed that they had to focus on the central fixation cross during stimuli presentation. Before the training session, participants performed ten test trials for each task (MR and TP respectively). As noted above, to avoid a possible bias favoring men (for reviews, Voyer, 2011), response time was unlimited and a feedback was given after each response. During the recording session, each trial of both tasks consisted in a sequential presentation of a fixation cross appearing at the center of the screen for 1000 ms, followed by a pair of stimuli presented until the participant pressed the response key (limited at 10 seconds), followed by a 1000 ms blank screen indicating the end of the trial. A total of 160 trials were presented, divided into 2 blocks of 80 trials for each of the tasks (MR and PT). The order of the tasks was counterbalanced across participants.

2.4 *Analysis of behavioral data*

Data were analyzed using SPSS software (SPSS 24, Chicago, USA). Trials yielding incorrect or missing responses were excluded from the analysis. Mean correct response time (FIG.2A) and number of correct responses (FIG.2B) were recorded during MR and PT tasks. Shapiro-Wilk test for normality was used to test the null-hypothesis that both the mean correct response times and number of correct responses were normally distributed. For the both the MR and PT tasks, the test failed to reject the null hypothesis ($p > .05$). Therefore, a one-way multivariate analysis of variance (MANOVA) was computed with mean correct response time and number of correct responses as dependent measures and sex (men, women) as between-subject measure. Subsequent univariate ANOVAs were used to determine how response time and number of correct responses differed across gender, and Bonferroni correction was applied to counteract

the problem of multiple comparisons.

INSERT FIGURE 2 ABOUT HERE

2.5 *EEG recording and ERP analysis*

Continuous EEG was recorded from 64 electrodes (Biosemi ActiveTwo system) placed according to the international 10-20 system location, as well as two pairs of bipolar electrodes recording the EOG in both vertical and horizontal directions. Two additional electrodes (active CMS: common mode sense and passive DRL: driven right leg) were used as reference and ground to compose a feedback loop for amplifier reference. All data were digitized at 1024 Hz and combined to stimulus delivery. Responses were recorded using E-Prime (Psychology Software Tools, Pittsburgh, Pa., USA). EEG signals were pre-processed with Brain Vision Analyzer (version 2.0.1.391; Brain Products GmbH). Data were first band-pass filtered offline between 1 Hz and 30 Hz using a zero-phase shift second-order Butterworth filter, and vertical and horizontal eye movement artifacts were corrected using an independent component analysis (ICA, Cardoso, 1998): for each subject, the two independent components representing vertical eye movements (and blinks) and horizontal eye movements (saccades) were eliminated and the corrected signal recomputed. The continuous recording was segmented into epochs from -100 ms pre-stimulus to 1000 ms post-stimulus and only epochs with correct responses were retained (corresponding to the 74% and 81% of the total number of epochs in the MR and PT tasks, respectively, see Table 1). Epochs were further excluded from ERP averaging if they contained activity exceeding $\pm$ 100 μV and channels containing excessive noise were replaced by linear spline interpolation of adjacent electrodes (Perrin, Pernier, Bertrand, Giard, & Echallier, 1987). This resulted in an average acceptance of 56.71 $\pm$ 7.67 trials for the MR task and 63.12 $\pm$ 6.74 trials for the PT task (corresponding to the ~71% and ~79% of the total number of epochs in

the MR and PT tasks, respectively). Data were then re-calculated to an average reference and individual artifact-free ERPs were computed for the MR and PT tasks. To investigate early attentional processes, brain dynamics were assessed by analyzing the ERPs from the onset until 450 ms after the stimulus. No baseline correction was applied.

INSERT TABLE 1 ABOUT HERE

2.5.1 *Consistency between subjects and topographical analysis of covariance*

A topographic consistency test (TCT), based on nonparametric randomization techniques, was applied to the ERPs of each experimental task to identify time periods where there was evidence for a consistent pattern of active sources (i.e., stable topographies) across subjects of each group (men, women). The TCT tests whether an event consistently activates brain electric sources across repeated measurements of event-related scalp field data (in our study these are single subject ERPs). This test is typically run at the beginning of the ERP analysis because, by identifying relevant time periods where there is evidence of a consistent relation between the event and the brain sources, false conclusions resulting from an incorrect selection of time windows of interest can be avoided (Koenig and Melie-Garcia, 2010). TCT was implemented on the open-source software RAGU (RAndomization Graphical User interface; Koenig et al., 2011; Habermann et al., 2018) based on Matlab (www.mathworks.com) and computed with 5'000 randomization runs and a *p* threshold of .05 (for a detailed description of the TCT see section SI.1 in the Supplementary information file). The time periods identified by TCT were retained for further analysis (TANCOVA, TANOVA, and microstate analysis).

A topographical analysis of covariance (TANCOVA; Koenig et al., 2008) was used to identify time periods where the ERP topographies in the MR and PT tasks significantly covaried with the accuracy of responses. To identify the ERP topographies that are proportional to the

accuracy of responses, the covariance map **V** (column vector of length equal to the number of electrodes, size [64 x 1] in this study) between the scalp potential **M** (matrix of size [subjects x electrodes], size [55 x 64] in this study) and the accuracy of responses **u** (column vector of length equal to the number of subjects, size [55 x 1] in this study) was computed at each time frame. Mathematically, the covariance map **V** was obtained, at each time frame, through the matrix product $\mathbf{V} = \mathbf{M^T}*\mathbf{u}$, where $\mathbf{M^T}$ is the transposition of **M**. The global field power (GFP) of the obtained covariance map (corresponding to the standard deviation of the covariance map values) was used as a global (across electrodes) measure of the size of the covariation between ERP potential and accuracy scores.

The statistical significance of this covariation was achieved by means of non-parametric randomization tests (Koenig et al., 2008): for each random iteration run, (i) the presumed covariance between scalp potential and accuracy of responses was removed by randomly multiplying $\mathbf{M^T}$ with a permuted version of accuracy of responses vector **u** (i.e., the vector elements are randomly permuted); (ii) the effect size (i.e., the GFP of this newly computed covariance map) was then calculated and retained. This procedure, repeated for 5'000 randomization runs, allowed estimating the distribution of our effect size under the null hypothesis of no covariation between scalp potentials and the accuracy of responses. Finally, the GFP of the covariance map obtained in the measured data was compared to the distribution of the GFP of the covariance maps obtained under the null hypothesis. The probability that the GFP of the covariance map obtained in the measured data was obtained by chance was defined as the percentage of observations where this latter was smaller than or equal to the GFP of the randomly obtained covariance maps. The TANCOVA was computed with RAGU using a *p* threshold of .05 (Koenig et al., 2011).

To protect the TANCOVA results from the possibility of obtaining false positives due to multiple testing, we estimated the probability of obtaining a certain number of time points

with a *p* value below .05 under the null hypothesis of no linear relation between ERP topographies and accuracy, and compared this estimate with the count of sub-threshold *p* values in our observed data. This was done to test whether the duration of a significant time period exceeded chance (Bermudez-Margaretto, 2015). This statistical test is available in RAGU and extensively described and detailed by Koenig et al. (2011) and Habermann et al. (2018) (for a detailed description of this procedure see section SI.3 in the Supplementary information file). After observing periods above duration threshold we computed the strength of the covariance maps in the ERP of each subject and evaluated the amount of explained variance: (i) first, the strength **s** (column vector of length equal to the number of subjects, size [55 x 1] in this study) of covariance map **V** in the ERP data **M** was computed (mathematically, this corresponds to the product **s = M*V)**; (ii) the percent common variance was then computed from the Pearson correlation coefficient between the strength of the covariance map **s** and the accuracy of responses **u**. For a comprehensive description of the procedure, see Koenig et al. (2008).

2.5.2 *Analysis of topographic differences and microstate analysis*

To supplement the TANCOVA described above and to clarify the gender impact on the accuracy of responses in each task, we used topographic analysis of variance (Lehman, 1987; Murray, et al., 2008; Koenig et al. 2011, TANOVA) and microstate analysis (Lehmann et al., 1987). Together, these analyses allowed us to determine the time periods of qualitative topographic differences between gender within each experimental task, and to provide a neurophysiological explanation for these differences.

Similar to the TANCOVA, the TANOVA is based on a non-parametric randomization test built on global dissimilarities between electric field topographies and does not require a preselection of time points for the analysis of the ERPs. Indeed, a key difference of this technique in comparison to a channel-wise comparison, in which any significant difference between experimental tasks or groups is detected at the level of the single electrode, is that the

TANOVA computes global dissimilarity of the whole electric field topographies between tasks or groups and test for significant topographic differences at each time point (for a detailed description of the TANOVA see section SI.2 in the Supplementary information file). The TANOVA was computed on the amplitude-normalized topographies (GFP=1), such that any significant topographic difference between gender could be attributed to partially different active sources of the evoked potentials, and not to differences in source strength (Murray et al., 2008; Koenig et al., 2011). In other words, this analysis allowed us to determine – for each experimental task- if and when the brain networks activated by men and women differ. The TANOVA was run using the open-source software RAGU for each experimental task (MR and PT respectively) with gender as between-subject factor, with 5'000 randomization runs, and a *p* threshold of .05 (Koenig et al., 2011; Ruggeri et al., 2019). As for the TANCOVA, to protect the TANOVA results from false positive caused by multiple testing, additional testing was performed to test whether the duration of significant time periods exceeded chance (for a detailed description of this procedure see section SI.3 in the Supplementary information file).

Although a significant effect revealed by the TANOVA is interpreted as the evidence of different topographies (or different configuration of brain sources if we consider the origin of the recorded scalp signal), this analysis alone does not unequivocally differentiate between the possible neurophysiological reasons for the observed differences (Murray et al., 2008). In other words, this analysis alone is unable to distinguish between topographic differences that arise from actual differences in map configurations, or from latency shifts of similar topographies between groups. In this regard, the comparison of microstate analysis results with significant effects revealed by the TANOVA offers the possibility to clarify the nature of the observed topographic differences between genders. This complementary approach allowed us to assess whether the observed differences between groups were associated with latency shifts of similar ERP topographies, or due to the presence of additional ERP topographies in one of

the two groups.

Dynamic variation in ERP topographies were thus further classified with microstate analysis. This technique reduces the ERP signal into a set of stable topographical map configurations, the microstate maps. The rationale behind this analysis is that stable brain functional states may vary between gender in their duration and/or strength of activation (Lehmann, et al., 1987; Lehmann, 1990; Murray, et al., 2008). The microstate analysis was computed using an algorithm implemented in RAGU (Koenig et al., 2010; Koenig et al., 2014). This procedure allows the identification of the optimal number of topographic map templates through a cross-validation procedure, an iterative technique that randomly splits the available ERPs into two subsets of the data (the learning and the test set, respectively), and then determines their topographical configurations. The number of topographic map templates is selected when the variance explained by the proposed templates reaches its maximum and does not increase with additional templates (for a detailed description of this procedure, see the SI file of Ruggeri et al., 2019). We applied this procedure 250 times, each time randomly partitioning the ERPs into training and learning sets, and testing between 1 and 10 microstate classes. To identify the topographical pattern of each microstate map, we used the k-means algorithm (Pascual-Marqui et al., 1995) with 50 random initializations.

Microstates analyses were computed separately for MR and PT tasks and values for onset, offset and duration of each microstate map were extracted from the fitting of the microstate maps on the grand mean ERPs of women and men. The onset and offset represent the beginning and the end time respectively of a given microstate, while the duration represents the difference between the offset and the onset time. For each experimental task, between groups differences of these parameters were estimated using randomization statistics with 5000 randomization runs (Koenig et al, 2010).

3. RESULTS

3.1 *Behavioral data*

Mean correct response times and mean accuracy scores (FIG2 and Table 1) recorded during MR and PT tasks were subjected to a one-way multivariate analysis of variance (MANOVA). The analysis revealed a statistically significant difference in performance based on gender (F (4, 50) = 7.309, p = < .0001; Wilk's Λ = .631, partial η^2 = .369). Subsequent univariate ANOVAs with Bonferroni correction revealed no gender effect in response time for either MR or PT, but a significant difference between men and women in correct responses for both the MR (Men: M = 62.00, SD = 5.26; Women: M = 54.61; SD = 6.20; F (1,53) = 22.796, Bonferroni-corrected p = < .001; η^2 = .301) and the PT task (Men: M = 67.24, SD = 5.553; Women: M = 62.27; SD = 5.02; F (1,53) = 12.020, Bonferroni-corrected p = .001; η^2 = .185).

3.2 *TCT and TANCOVA*

The TCT test was applied on the pre-processed ERPs of the PT and MR tasks and showed consistent topographies across subjects for each group, over the period considered (from 0 to 450 ms). The TANCOVA analysis was thus performed over the entire period to evaluate the linear association between ERP topographies and accuracy scores. The TANCOVA analysis conducted in the PT task revealed covarying ERP topographies with p value below the .05 threshold in the time periods ranging from 164.6 to 177.3 ms and from 403 to 411 ms. However, these short periods of significance did not survive the correction for multiple testing and were not retained for further investigation.

In the MR task, the TANCOVA analysis revealed significant ERP topographic covariances in the time periods from 113.9 to 144.1 ms and from 167.6 to 195.9 ms (FIG. 3A), corroborated by multiple testing correction performed between 0 and 300 ms. The topography of the covariance map observed between 113.9 and 144.1 ms (FIG. 3B, top) indicated that the number of correct responses covaried with positive potential at anterior sites and negative potential at posterior and occipital sites. A linear regression analysis performed between the

strength of this covariance map in the ERP data and the number of correct responses revealed that it explained 17% of the variance (FIG. 3B, bottom; $r = .41$, $p = .002$). More specifically, the results showed that, within this period, this topography was less present among individuals with high number of correct responses, but more strongly present in those with lower performance rates. The topography of the covariance map observed during the second time period (FIG. 3C, top) indicated that the number of correct responses covaried with positive potential at bilateral posterior-occipital sites and negative potential at central-parietal sites. A linear regression analysis performed between the strength of this covariance map in the ERP data and the number of correct responses revealed that an amount of variance of 18% was explained (FIG. 3C, bottom; $r = .43$, $p = .005$). More specifically, the results showed that, within this period, this topography was more strongly present among individuals with higher number of correct responses, but less present in those with low accuracy scores.

INSERT FIGURE 3 ABOUT HERE

3.3 *Topographical and microstate analysis*

Topographical differences between genders for each experimental task (MR and PT respectively) were examined using the TANOVA and microstate analysis.

In the PT task, the TANOVA revealed significant differences between topographies of women and men in the time periods from 103 to 138 ms and from 210 to 252 ms (FIG. 4A), corroborated by multiple testing correction performed between 0 and 300 ms. Qualitative examination of the topographies of both groups (Fig. 4B) during the time period between 103 and 138 ms showed a difference in the time course of the topography characterized by bilateral occipital positivity and central negativity that seemed to last longer in men than in women. Congruently to this observation, post-hoc channel wise *t*-tests (*t*-map contrasts) between the

average topographies of men and women in this time window (FIG. 4B) was characterized by a more positive potential over posterior and occipital electrodes and a more diffused negative potential localized over central-frontal regions (p = .003; t_{max} = 3.83 at electrode PO8; t_{min} = -5.26 at electrode FC3). Between 210 and 352 ms, the t-map contrasts between the average topographies of men and women revealed a more diffused positive potential over central-frontal electrodes and a more negative potential localized over posterior-occipital regions (p = .003; t_{max} = 3.3 at electrode C3; t_{min} = -3.45 at electrode P8). Microstate analysis indicated five microstate maps accounting for 90% of the ERP variance. The microstate maps, together with their fit with the grand-mean ERPs of women and men, are shown in FIG. 4C. The statistics computed on the onset, duration and offset of the microstate maps revealed significant differences between women and men for microstate maps 2 and 5 (Table 2): microstate map 2, which ended later (p = .003) in men (140.2 ms) compared to women (116.8 ms), lasted longer (p = .012) in men (58.6 ms) compared to women (37.1 ms); microstate map 5 occurred earlier (p = .049) in women (224.2 ms) compared to men (243.7). The time periods in which these time shifts occurred overlapped with the time windows of significant topographic differences previously outlined by the TANOVA analysis. More specifically, the topographic differences between gender observed in the time periods from 103 to 138 ms and from 210 to 252 ms were compatible with effects due to the longer duration of microstate map 2 and the later onset of microstate map 5 in men compared to women.

INSERT TABLE 2 ABOUT HERE

INSERT FIGURE 4 ABOUT HERE

In the MR task, the TANOVA revealed significant differences between topographies of women and men between 119 and 176 ms (FIG. 5A), corroborated by multiple testing correction performed between 0 and 300 ms. Within this time window, the topographies of women seemed to shift more rapidly from a scalp topography characterized by positive potential at bilateral posterior-occipital and negative potential at central sites to a configuration displaying positive potential at central-frontal sites and negative potential over posterior and occipital sites (FIG. 5B). Post-hoc channel-wise t-test within this time period contributed to a better understanding of the topographic distribution of the observed differences between women and men. The t-map contrasts between the average topographies of men and women (FIG. 5B) revealed that the topography of men was characterized by a more positive potential over bilateral posterior and occipital regions and a more negative potential over central-frontal electrodes ($p = .001$; $t_{max} = 4.46$ at electrode P8; $t_{min} = -5.97$ at electrode FC1). The microstate analysis revealed an optimum of five microstate maps accounting for ~80% of the ERP variance. The fitting of the microstate maps to the grand-mean ERPs resulted in the same five microstate maps occurring in the same temporal sequence in both women and men (FIG. 5C). Randomization-based statistical tests on the microstate map parameters onset, offset and duration (Table 3) revealed significant temporal shifts between topographies of women and men inducing offsets in microstates map 2 and 3. Precisely, they revealed an earlier end for microstate map 2 ($p = .008$) in women (143.1 ms) compared to men (156.8 ms), and an onset of microstate map 3 occurring earlier ($p = .041$) in women (149 ms) compared to men (164.6 ms). No significant differences were observed in the parameters of microstate maps 1, 4 and 5. The time period covered by microstate maps 2 and 3 during which significant latency shifts were observed between women and men, also corresponded to the time period (between 119 and 176 ms) where topographic differences between gender occurred (FIG. 5A and 5B). This similarity suggested that the topographic differences revealed by the TANOVA could be

explained by altered latencies of the underlying ERP components between women and men. To better illustrate this aspect, ERPs dynamics and variability across men and women in the MR task are shown at frontal-central, left posterior-occipital and right-posterior occipital ROIs (supplementary Figure S1A, S1B and S1C, respectively). Also shown is the time period of significant differences between women and men as revealed by the TANOVA, and the time periods of appearance of the microstate maps 2 and 3 whose onset and offset significantly differed between women and men.

INSERT TABLE 3 ABOUT HERE

INSERT FIGURE 5 ABOUT HERE

Since the time period of observed topographical differences revealed by the TANOVA and the microstate analysis in the MR task overlapped with the time period of significant covariation of ERP topographies and number of correct responses revealed by the TANCOVA in the same task, an additional analysis was run to quantify whether the latency shifts of microstate maps 2 and 3 between women and men could be associated with the covariance maps outlined in the TANCOVA analysis (FIG. 3B and 3C). This aspect was investigated by examining the spatial correlation between microstate maps 2 and 3 and the covariance maps described in Section 3.2 (FIG 6). The results revealed a positive correlation between microstate map 3 and the covariance map observed in the time periods ranging from 113.9 to 144.1 ms (r = .85) and between microstate map 2 and the covariance map observed in the time periods ranging from 167.6 to 195.9 ms (r = .86).

INSERT FIGURE 6 ABOUT HERE

4. Discussion

The current work used the full spatial and temporal information of EEG to examine potential gender differences – regarded here as a typical example of interindividual differences – in mental rotation tasks. It aimed at evaluating the hypothesis that interindividual performance differences, measured in mental rotation tasks, depend on the length of transient brain states allocates to the early stages of visual processing.

Behaviorally, MR and PT tasks reflected the male advantage habitually observed in tasks involving visuospatial mental transformations (Voyer, Voyer & Bryden, 1995; Kaiser, et al., 2008). Compared to women, men's accuracy was significantly higher in both mental rotation and perspective tasks despite an unlimited response time. As expected, since time was not limited, there was no significant gender differences in reaction time.

Interestingly, our novel approach of data analysis brought to light new findings regarding interindividual differences observed in mental rotation tasks. The main findings were found within a time period ranging from 114 to 196 ms post stimulus, coinciding with P1/N1 components described in the literature. In this time window, we showed an association between early attentional processes and accuracy scores, revealed by a close association between spatial covariance and microstate topographies, exhibiting substantial genders differences. As we did not find similar studies in PT, our discussion will focus on MR results with ad hoc comments on PT tasks.

As expected, we found evidence for consistent patterns of active sources within the time range considered (from 0 to 450 ms) in both tasks. One of the most exciting results was the observation that early potentials, regarded as reflecting early visual attention process, predicted score accuracy in the classical mental rotation task. In the PT task, we showed an association in a short time period ranging from 164 to 177 ms but, perhaps due to the short time duration,

the statistical significance did not survive to the correction for multiple testing. Thus, the temporal covariation between the global scalp field potentials and the accuracy was not assessed. In the MR task, ERP topographic covariances were observed in the time periods ranging from 114 to 144 ms and from 168 to 200 ms with maximal covariance map values at the central posterior sites and at bilateral parietal/occipital sites. Both the time windows and topographies associated with score accuracy were in line with earlier ERP component studies reflecting attentional processes (Mangun, & Hillyard, 1988; Mangun, & Hillyard, 1991; Luck, et al., 1990; Luck, Fan, & Hillyard, 1993; Gomez Gonzalez, 1994; Mangun, 1995). These studies identified changes in the amplitude of P1/N1 ERP components mostly characterized by distributions over the posterior brain regions (parietal, temporal, and occipital) although, depending on the location of the stimulus and the type of task, N1 could be distributed in frontal regions (Vaquero, et al., 2004).

As our results confirmed the existence of a relationship between early visual attention processes and score accuracy, additional analyses were performed in each task for the two genders. As expected, these revealed significant topographic differences between women and men in both PT and MR tasks within the temporal period concordant with P1/N1 ERP components. In the PT task, although microstate sequences did not differ across gender, a consistent shift was observed for microstates 2 and 5. The duration of microstate 2 was significantly longer in men compared to women, and ended significantly later in men than women. The onset of microstate 5 in women occurred significantly earlier in comparison to men. An examination of the time period where significant topographic differences occurred revealed a bilateral positivity lasting longer in men than in women, a more centro-parietal distribution in men compared with women, as well as a more occipital distribution in women compared with men. As with the PT task, microstate sequences did not differ across gender in the MR task, but a consistent shift was observed for microstates 2 and 3. The onset of microstate

2 did not differ between men and women but ended significantly later in the former group. Moreover, the onset of microstate 3 in women occurred significantly earlier compared to men. An examination of the time period during which significant topographic difference occurred revealed a more rapid shift from a bilateral positive potential at posterior-occipital sites and negative potential over central-frontal regions, to a negative potential at central-frontal sites and positive potential over posterior-occipital regions in women compared with men. The high consistency between the results of covariance analysis and microstates within each sex suggests that the difference in accuracy in MR between women and men is, at least in part, based on the processing of early visual information.

In a classic mental rotation task, the observation of a greater recruitment in parietal and temporal lobes in males and in prefrontal regions in women were interpreted as a possible mark of a more effortful maintenance of visuospatial transformations in working memory in women (Butler, et al., 2006). Although appealing, such an interpretation remains questionable because of the limited temporal resolution and the slow response of the BOLD signal in fMRI techniques. Indeed, these approaches do not allow to determine whether an increased activation results from higher neuronal activity per time unit, or from the same degree of processing but for a longer time. Surprisingly, most of the studies on mental rotation were mostly focused on later, higher order cognitive processes, leaving aside more basic aspects as visual perception or attention that might dramatically impact later processes. The few ERP studies exploring gender differences at the early processing stage of mental rotation tasks yielded mixed results. They showed either an absence of differences in the P1/N1 amplitude (Best, Heil, & Konrad, 2010), or a greater amplitude in women compared to men (Gootjes, 2008), or still longer P1 latencies in men, along with lower N1 amplitudes, compared with women (Jaušovec, 2012).

Contrary to these studies based on conventional ERP analyses performed at specific, a priori-selected electrode sites, our results were based on an unbiased measure (GFP) of the

electrical signal at each time point allowing the exploration of the brain dynamics underlying gender differences on a millisecond scale. At the scalp level, the microstate and TANOVA analysis were inconsistent with the idea that additional microstate maps might explain gender differences. In contrast, these analyses demonstrated that the electrical dynamics differed in duration at the early steps of information processing, as shown by the increased duration of microstate map 2 in men compared to women.

Taken together, these findings yield new evidence supporting our hypothesis that interindividual performance differences in mental rotation tasks depend on the length of transient brain states allocates to the early stages of visual processing. Although the two mental rotation tasks used here cannot be directly compared, they both suggest that the male brain engages more time processing visual spatial information resulting in a longer bilateral positive potential at posterior-occipital sites. In return, the female brain initiates an earlier processing of non-spatial stimuli information resulting in a faster transition from a bilateral positive potential of posterior-occipital sites to a negative potential at central-frontal sites. This interpretation is consistent with one of the most rigorous ERP studies that hypothesized that the processing of nonspatial features reflected in N2 component is highly dependent on the prior selection for location, reflected in the P1 and N1 components (Hillyard and Anllo-Vento,1998). It should be noted that our hypothesis that interindividual differences may stem from differences in the time devoted to the early processing of visual spatial stimuli does not preclude the existence of possible differences in working memory or other higher functions. Indeed, if more time is devoted for processing visual spatial stimuli, it is conceivable that a larger amount of information is needed to be stored in memory for later use. However, we cannot exclude that some brain mechanisms might promote a better compression of the chain of information, eliminating needless redundancy, thus reducing the necessity for a larger storage system. Only further research might disentangle such issues.

To conclude, a more complete utilization of the spatiotemporal information contained in EEG recordings can provide important insights about the impact of early visual processes on interindividual differences, particularly across gender, and thus shed new light on differences in cognitive strategies.

REFERENCES

Beste, C., Heil, M., & Konrad, C. (2010). Individual differences in ERPs during mental rotation of characters: lateralization, and performance level. *Brain and cognition*, *72*(2), 238–243. https://doi.org/10.1016/j.bandc.2009.09.005

Bermúdez-Margaretto, B., Beltrán, D., Domínguez, A., & Cuetos, F. (2015). Repeated Exposure to "meaningless" Pseudowords Modulates LPC, but Not N(FN)400. *Brain topography*, 28(6), 838–851. https://doi.org/10.1007/s10548-014-0403-5

Blough, P. M., & Slavin, L. K. (1987). Reaction time assessments of gender differences in visual-spatial performance. *Perception & psychophysics*, 41(3), 276–281. https://doi.org/10.3758/bf03208225

Butler, T., Imperato-McGinley, J., Pan, H., Voyer, D., Cordero, J., Zhu, Y. S., Stern, E., & Silbersweig, D. (2006). Sex differences in mental rotation: top-down versus bottom-up processing. *NeuroImage*, 32(1),445–456. https://doi.org/10.1016/j.neuroimage.2006.03.030

Cardoso, J. F. (1998). Blind signal separation: statistical principles. *Proceedings of the IEEE*, 86(10), 2009-2025. https://doi.org/10.1109/5.720250.

Cooke-Simpson, A., & Voyer, D. (2007). Confidence and gender differences on the Mental Rotations Test. *Learning and Individual Differences*, 17(2), 181-186. https://doi.org/10.1016/j.lindif.2007.03.009.

Debelak, R., Gittler, G., & Arendasy, M. (2014). On gender differences in mental rotation processing speed. *Learning and Individual Differences*, 29, 8-17. https://doi.org/10.1016/j.lindif.2013.10.003.

Delgado, A. R., & Prieto, G. (1996). Sex differences in visuospatial ability: do performance factors play such an important role?. *Memory & cognition*, 24(4), 504–510. https://doi.org/10.3758/bf03200938

Feng, Q., Zheng, Y., Zhang, X., Song, Y., Luo, Y. J., Li, Y., & Talhelm, T. (2011). Gender differences in visual reflexive attention shifting: evidence from an ERP study. Brain research, 1401, 59–65. https://doi.org/10.1016/j.brainres.2011.05.041

Goldstein, D., Haldane, D., & Mitchell, C. (1990). Sex differences in visual-spatial ability: the role of performance factors. Memory & cognition, 18(5), 546–550. https://doi.org/10.3758/bf03198487

Gomez Gonzalez, C. M., Clark, V. P., Fan, S., Luck, S. J., & Hillyard, S. A. (1994). Sources of attention-sensitive visual event-related potentials. *Brain topography*, 7(1), 41–51. https://doi.org/10.1007/BF01184836

Gootjes, L., Bruggeling, E. C., Magnée, T., & Van Strien, J. W. (2008). Sex differences in the latency of the late event-related potential mental rotation effect. *Neuroreport*, 19(3), 349–353. https://doi.org/10.1097/WNR.0b013e3282f519b3

Habermann, M., Weusmann, D., Stein, M., & Koenig, T. (2018). A Student's Guide to Randomization Statistics for Multichannel Event-Related Potentials Using Ragu. *Frontiers in neuroscience*, 12, 355. https://doi.org/10.3389/fnins.2018.00355

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the United States of America,* 95(3), 781–787. https://doi.org/10.1073/pnas.95.3.781

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral cortex* (New York, N.Y.: 1991), 14(2), 132–142. https://doi.org/10.1093/cercor/bhg111

Jaušovec N. (2012). Sex differences in event-related potential components during the solution of complex mental rotation tasks. *Neuroreport*, 23(6), 360–363. https://doi.org/10.1097/WNR.0b013e328351dd9a

Jausovec, N., & Jausovec, K. (2009). Do women see things differently than men do?. *NeuroImage*, 45(1), 198–207. https://doi.org/10.1016/j.neuroimage.2008.11.013

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology,* 8(4), 441-480. https://doi.org/10.1016/0010-0285(76)90015-3.

Kaiser, S., Walther, S., Nennig, E., Kronmüller, K., Mundt, C., Weisbrod, M., ... & Vogeley, K. (2008). Gender-specific strategy use and neural correlates in a spatial perspective taking task. *Neuropsychologia*, 46(10), 2524–2531. https://doi.org/10.1016/j.neuropsychologia.2008.04.013

Koenig, T., & Melie-García, L. (2010). A method to determine the presence of averaged event-related fields using randomization tests. *Brain topography*, 23(3), 233–242. https://doi.org/10.1007/s10548-010-0142-1

Koenig, T., Kottlow, M., Stein, M., & Melie-García, L. (2011). Ragu: a free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Computational intelligence and neuroscience*, 2011, 938925. https://doi.org/10.1155/2011/938925

Koenig, T., Melie-García, L., Stein, M., Strik, W., & Lehmann, C. (2008). Establishing correlations of scalp field maps with other experimental variables using covariance analysis and resampling methods. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology,* 119(6), 1262–1270. https://doi.org/10.1016/j.clinph.2007.12.023

Koenig, T., Stein, M., Grieder, M., & Kottlow, M. (2014). A tutorial on data-driven methods for statistically assessing ERP topographies. *Brain topography*, 27(1), 72–83. https://doi.org/10.1007/s10548-013-0310-1

Lehmann, D. (1987). Principles of spatial analysis. In A. Gevins & A. Rémond (Eds.), *Handbook of electroencephalography and clinical neurophysiology: Methods of analysis of brain electrical and magnetic signals* (Vol.1, pp 309-354). Amsterdam, Netherlands: Elsevier.

Lehmann, D. (1990). Brain Electric Microstates and Cognition: The Atoms of Thought. In: John E.R., Harmony T., Prichep L.S., Valdés-Sosa M., Valdés-Sosa P.A. (Eds.), *Machinery of the Mind*. Boston, MA: Birkhäuser.

Lehmann, D., Ozaki, H., & Pal, I. (1987). EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and clinical neurophysiology*, 67(3), 271–288. https://doi.org/10.1016/0013-4694(87)90025-3

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and clinical neurophysiology*, 48(6), 609–621. https://doi.org/10.1016/0013-4694(80)90419-8

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in cognitive sciences*, 4(11), 432–440. https://doi.org/10.1016/s1364-6613(00)01545-x

Luck, S. J., & Girelli, M. (1998). Electrophysiological approaches to the study of selective attention in the human brain. In R. Parasuraman (Ed.), *The attentive brain* (p. 71–94). The MIT Press

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences of the United States of America,* 95(3), 825–830. https://doi.org/10.1073/pnas.95.3.825

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. Journal of experimental psychology. *Human perception and performance*, 20(5), 1000–1014. https://doi.org/10.1037//0096-1523.20.5.1000

Luck, S. J., Fan, S., & Hillyard, S. A. (1993). Attention-related modulation of sensory-evoked brain activity in a visual search task. *Journal of cognitive neuroscience*, 5(2), 188–195. https://doi.org/10.1162/jocn.1993.5.2.188

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and clinical neurophysiology*, 75(6), 528–542. https://doi.org/10.1016/0013-4694(90)90139-b

Mangun G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32(1), 4–18. https://doi.org/10.1111/j.1469-8986.1995.tb03400.x

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. Journal of experimental psychology. *Human perception and performance*, 17(4), 1057–1074. https://doi.org/10.1037//0096-1523.17.4.1057

Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual attention: behavioral and electrophysiological evidence. *Electroencephalography and clinical neurophysiology*, 70(5), 417–428. https://doi.org/10.1016/0013-4694(88)90019-3

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain topography*, 20(4), 249–264. https://doi.org/10.1007/s10548-008-0054-5

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE transactions on bio-medical engineering*, 42(7), 658–665. https://doi.org/10.1109/10.391164

Perrin, F., Pernier, J., Bertrand, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and clinical neurophysiology*, 66(1), 75–81. https://doi.org/10.1016/0013-4694(87)90141-6

Piaget, J., & Inhelder, B. (1963). *The child's conception of space*. London: Routledge & KeganPaul.

Pletzer B. (2014). Sex-specific strategy use and global-local processing: a perspective toward integrating sex differences in cognition. Frontiers in neuroscience, 8, 425. https://doi.org/10.3389/fnins.2014.00425

Pylyshyn Z. W. (1979). The rate of "mental rotation" of images: a test of a holistic analogue hypothesis. *Memory & cognition*, 7(1), 19–28. https://doi.org/10.3758/bf03196930

Ruggeri, P., Meziane, H. B., Koenig, T., & Brandner, C. (2019). A fine-grained time course investigation of brain dynamics during conflict monitoring. *Scientific reports*, 9(1), 3667. https://doi.org/10.1038/s41598-019-40277-3

Shearer, D. E., Cohn, N. B., Dustman, R. E., & LaMarche, J. A. (1984). Electrophysiological correlates of gender differences: A Review. *American Journal of EEG Technology*, 24(2), 95–107. https://doi.org/10.1080/00029238.1984.11080133

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* (New York, N.Y.), 171(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Taylor M. J. (2002). Non-spatial attentional effects on P1. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, 113(12), 1903–1908. https://doi.org/10.1016/s1388-2457(02)00309-7

Vaquero, E., Cardoso, M. J., Vázquez, M., & Gómez, C. M. (2004). Gender differences in event-related potentials during visual-spatial attention. *The International journal of neuroscience*, 114(4), 541–557. https://doi.org/10.1080/00207450490422056

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological bulletin*, 117(2), 250–270. https://doi.org/10.1037/0033-2909.117.2.250

Voyer D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: a meta-analysis. *Psychonomic bulletin & review,* 18(2), 267–277. https://doi.org/10.3758/s13423-010-0042-0

Yuan, J., He, Y., Qinglin, Z., Chen, A., & Li, H. (2008). Gender differences in behavioral inhibitory control: ERP evidence from a two-choice oddball task. *Psychophysiology*, 45(6), 986–993. https://doi.org/10.1111/j.1469-8986.2008.00693.x

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation*, 2(4), 315-332.

Zani, A., & Proverbio, A. M. (2012). Is that a belt or a snake? Object attentional selection affects the early stages of visual sensory processing. *Behavioral and brain functions : BBF*, 8, 6. https://doi.org/10.1186/1744-9081-8-6

Zani, A., Marsili, G., Senerchia, A., Orlandi, A., Citron, F. M., Rizzi, E., & Proverbio, A. M. (2015). ERP signs of categorical and supra-categorical processing of visual information. *Biological psychology*, 104, 90–107. https://doi.org/10.1016/j.biopsycho.2014.11.012

TABLES

	MR		PT	
	Correct response	RT	Correct response	RT
Women	54.61 ± 6.20	3397.80 ± 940.61	62.27 ± 5.02	3049.09 ± 896,72
Men	62.00 ± 5.26	3895.35 ± 740.81	67.24 ± 5.55	3065.44 ± 802,02

Table 1: Mean and standard deviation of correct responses and mean response time by sex and task.

Microstate	Parameter	Women	Men	p-value
Map 1	Onset [ms]	0	0	ns
	Duration [ms]	121.1	105.5	ns
	Offset [ms]	158.8	164.6	ns
Map 2	Onset [ms]	79.7	81.6	ns
	Duration [ms]	37.1	58.6	$p = .012$
	Offset [ms]	116.8	140.2	$p = .003$
Map 3	Onset [ms]	158.8	165.6	ns
	Duration [ms]	33.2	36.1	ns
	Offset [ms]	192	201.7	ns

Microstate	Parameter	Women	Men	p-value
Map 4	Onset [ms]	192	201.7	ns
	Duration [ms]	87.9	106.4	ns
	Offset [ms]	375.6	386.3	ns
Map 5	Onset [ms]	224.2	243.7	$p = .049$
	Duration [ms]	170.7	143.4	ns
	Offset [ms]	450	450	ns

Table 2: Onsets, offsets and durations (in milliseconds) of microstate maps fitted on the ERPs of women and men in the PT condition. Also shown are the *p*-values for the group effect resulting from the microstate parameters statistical analysis. "ns": not significant.

Microstate	Parameter	Women	Men	p-value
Map 1	Onset [ms]	0	0	ns
	Duration [ms]	78.7	85.5	ns
	Offset [ms]	78.7	85.5	ns
Map 2	Onset [ms]	78.7	85.5	ns
	Duration [ms]	65.4	72.3	ns
	Offset [ms]	143.1	156.8	$p = .008$
Map 3	Onset [ms]	149	164.6	$p = .041$
	Duration [ms]	53.7	44	ns
	Offset [ms]	201.7	208.6	ns
Map 4	Onset [ms]	201.7	208.6	ns
	Duration [ms]	82.1	78.1	ns
	Offset [ms]	283.8	286.7	ns
Map 5	Onset [ms]	283.7	287.7	ns
	Duration [ms]	166.3	162.3	ns
	Offset [ms]	450	450	ns

Table 3: Onsets, offsets and durations (in milliseconds) of microstate maps fitted on the ERPs of women and men in the MR condition. Also shown are the *p*-values for the group effect resulting from the microstate parameters statistical analysis. "ns": not significant.

FIGURE CAPTIONS

Figure 1: (A) In the MR condition participants had to judge whether the images on the right were the same (top) or different (bottom), (B) while in the PT condition they had to judge whether the images on the right coincided (top) or not (bottom) with the location of an observer, symbolized by an eye (arrow).

Figure 2: Violin plot representation of the (A) response time and the (B) number of correct responses distributions for the MR and PT tasks. The light blue dots represent the original data and the violin plot outlines represent kernel probability densities of the original data. The white dot corresponds to the median value of the distribution, while the dark grey line illustrates the inter-quartile range between the 25th and the 75th percentile.

Figure 3: Results of the TANCOVA analysis performed using the ERPs and the accuracy scores of the MR condition. (A) The *p*-value of the TANCOVA is plotted as a function of time. The grey shadowed areas highlight the time periods of significant covariation of ERP topographies and accuracy scores. (B and C) The mean topographies of the covariance maps observed from 113.9 to 144.1 ms and from 167 to 195.9 ms (top) are illustrated. The graphs below indicate the results of the linear regression analysis performed between their observed strengths in the ERPs and the accuracy scores.

Figure 4: TANOVA and microstate analysis results outlining periods of topographic differences of the grand-mean ERPs of the PT condition between women and men. (A) The *p*-value of the TANOVA analysis is plotted as a function of time. The grey shadowed areas highlight the time periods of significant differences between topographies of women and men. (B) Time course evolution of the ERP topographies of women (top) and men (bottom). Topographies within the time period of significant topographic difference are illustrated in gray boxes and displayed at higher time resolution (topographies' color bar is scaled between -3 and +3 μV); also shown are standard *t*-maps contrasting average topographies of men and women in the same time windows (*t*-map' color bar is scaled between -5 and +5). Positive (red) and negative (blue) *t*-values indicate more positive and more negative potential for men compared to women, respectively. (C) Topographies are shown for the five microstate maps obtained from the cross-validation procedure, along with their periods of appearance (coloured areas) displayed under the GFP (vertical axis) in the grand mean data of women (top) and men (bottom) as a function of time (horizontal axis).

Figure 5: TANOVA and microstate analysis results outlining periods of topographic differences of the grand-mean ERPs of the MR condition between women and men. (A) The *p*-value of the TANOVA analysis is plotted as a function of time. The grey shadowed area highlights the time period of significant differences between topographies of women and men. (B) Time course evolution of the ERP topographies of women (top) and men (bottom). Topographies within the time period of significant topographic difference are shown in the gray box and displayed at higher time resolution (topographies' color bar is scaled between -3 and +3 μV); also shown is the standard *t*-map contrasting average topographies of men and women in the same time window (*t*-map' color bar is scaled between -5 and +5). Positive (red) and negative (blue) *t*-values indicate more positive and more negative potential for men compared to women,

respectively. (C) Topographies of the five microstate maps obtained from the cross-validation procedure, along with their periods of appearance (coloured areas) displayed under the GFP (vertical axis) in the grand mean data of women (top) and men (bottom) as a function of time (horizontal axis).

Figure 6: Spatial correlation between the microstate maps 2 and 3 identified by the microstate analysis and the covariance maps revealed by the TANCOVA analysis in the MR condition.

Supplementary Information

Interindividual differences in brain dynamics of early visual processes: impact on score accuracy in mental rotation task

Paolo Ruggeri[a], Nam Nguyen[a], Alan J. Pegna[b] and Catherine Brandner[a]*

[a] Brain Electrophysiology Attention Movement Laboratory, Institute of Psychology, University of Lausanne, Switzerland.

[b] School of Psychology, The University of Queensland, Brisbane, Qld-4072, Australia.

* Corresponding author: Catherine.Brandner@unil.ch

SI METHODS

SI.1 Topographic consistency test

The TCT uses the GFP of the grand-average ERP across subjects as a measure of effect size (i.e., a measure of consistency of the topography across subjects). The assumption behind TCT is that the GFP of the grand-average ERP across subjects at a specific moment in time depends on both the GFP of the individual ERP and on the consistency of the topography over subjects. For instance, a "low" -across subject- topography consistency will display a GFP of the grand-average inconsistently lower than the mean of the individual GFP values. Instead, a "high" -across subjects -topography consistency will lead to a grand-average GFP slightly lower than the mean of the individual GFP values.

When we apply TCT to a specific dataset, we test whether a given effect size could have been observed by chance. This is done by destroying the assumed structure of the data by shuffling the

measured potentials values across electrodes in each individual ERP map. This procedure smartly preserves the GFP of the individual map (i.e., the standard deviation of the potential values is not dependent on the location of the measured values) but destroys the consistency of the topography across subjects. The null hypothesis that is tested with TCT is thus that the GFP of the grand-average ERP across subjects before shuffling is about equally large as after random permutation. The null hypothesis can be rejected if the GFP of the grand-average ERP across subjects is consistently larger when the original data structure is preserved as opposed to when the location of measured potential values across the scalp is shuffled. In other words, the probability of the null hypothesis is defined as the number of randomizations runs yielding a GFP larger than or equal to the GFP obtained with the original data structure (i.e., correct channel order). In the case of an ERP analysis, this procedure is repeated for each ERP time point and separately for each experimental group (or conditions), resulting in a moment by moment test of topographic consistency.

SI.2 Topographic analysis of variance

The TANOVA is used in this study to test for significant topographic differences between gender. The TANOVA is a randomization test that uses the strength of the topographic difference between groups topographies as a measure of effect size (i.e., the GFP of the difference topographies). First, the scalp field topographies are averaged separately for both groups yielding two group-wise grand-average topographies. The GFP of the difference between the group-wise topographies is used as indicator of the strength of the difference and serves as effect size's measure. To create instances of effect size under the null hypothesis, the underlying structure of the data is destroyed by randomly permuting the ERP data of each subject between groups. After, the (random) group specific grand-average ERPs is computed, and the GFP of the topography obtained from the difference between the two (random) group-wise topographies is computed again. This represents one instance of the effect size under the null hypothesis. This procedure, repeated for many random runs (i.e., by randomly permuting the ERP data of each subject between groups), allows estimating the distribution of the

effect size under the null-hypothesis. To estimate how the observed topographic difference between groups compare with the alternative hypothesis of no topographic differences between groups, the GFP of the difference map obtained in the measured data is compared to the distribution of the GFP of the difference maps obtained under the null hypothesis. The probability that the observed GFP was obtained by chance is defined as the percentage of observations where this GFP was smaller than or equal to the GFP of the randomly obtained topographic differences. Such a test can be applied to each time point of the ERP trace and the duration of the significant differences validated with a global duration statistics test (see next session, SI.3).

SI.3 Global duration statistics

Tests such as the TANCOVA and TANOVA that are typically performed at each time-point of the ERP trace, require methods of accounting for multiple testing over time. Indeed, as previously described for the TANOVA (this is valid also for the TANCOVA), a test for significance can be computed for every time point, yielding a distribution of p-values: a subset of these will be below a chosen threshold by chance, and will thus constitute false positive if this is the only criterion for significance. One approach that can be used to minimize this problem is to quantify the duration of contiguous periods with sub-threshold p-values. To estimate the probability that a given duration of sub-threshold p-values is present under the null-hypothesis, the results obtained from the randomization procedure are used again in this test. For every randomization run computed during the TANOVA (or the TANCOVA), p-values can be computed by comparing the random differences with those obtained in all other randomization runs. This allows quantifying the duration of periods with sub-threshold p-values that are expected to be observed under the null-hypothesis, yielding a distribution of the duration of false-positives under the null-hypothesis. The duration of contiguous sub-threshold p-values in the measured data can then be compared to this distribution (i.e., the distribution of contiguous false positives' duration in the randomized data). A 5% p-threshold for the overall significance indicates the duration of periods of contiguous p-values that is larger than 95%

of the false positive duration of contiguous *p*-values obtained in the random data (thus producing an

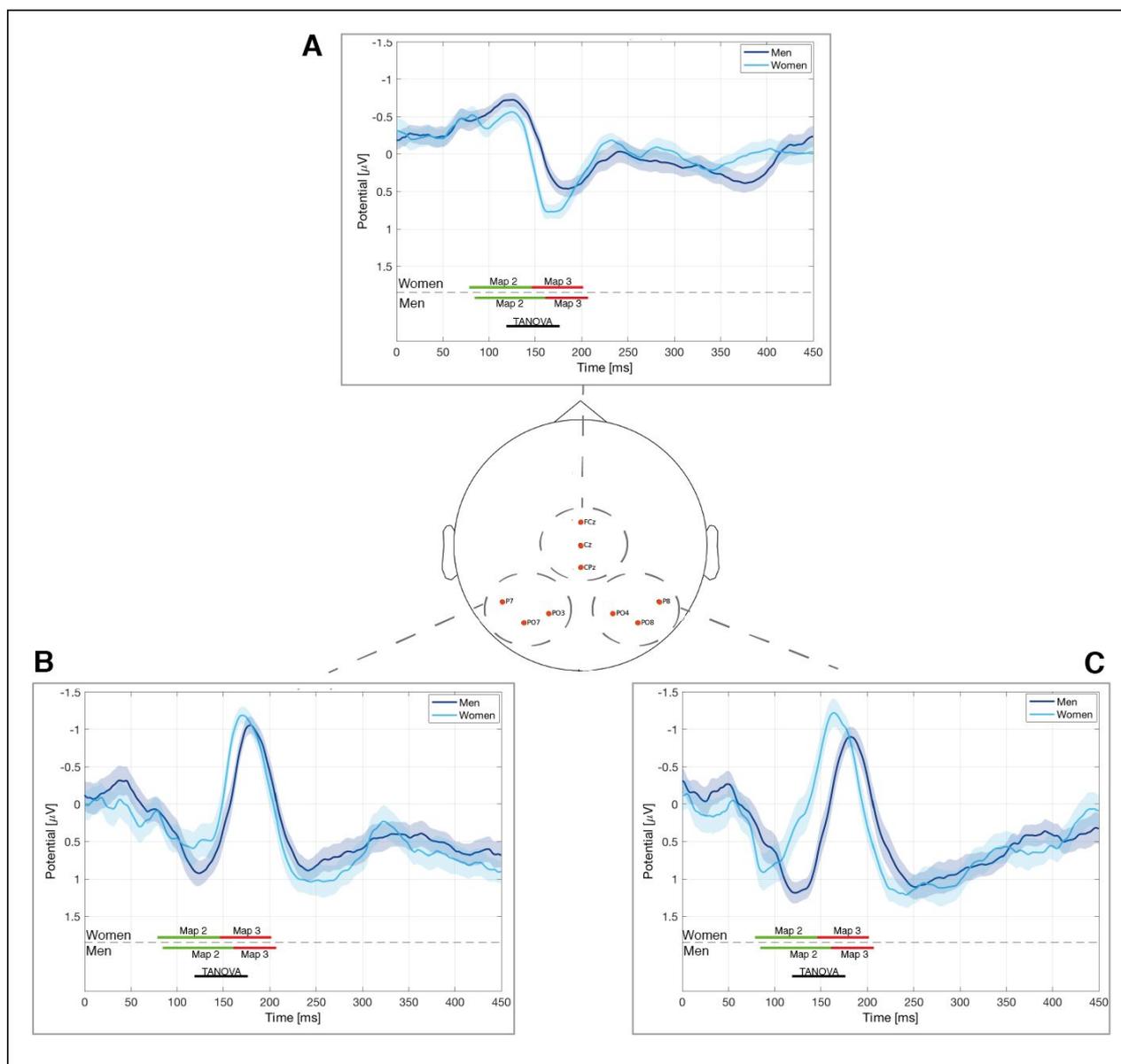overall 5% false positive rate)

SI FIGURES

Fig. S1: ERP dynamics at a (A) frontal-central ROI, (B) left posterior-occipital ROI, and (C) right posterior-occipital ROI in the MR condition for women and men (continuous ERP lines). The shadowed areas represent the surface that encloses the Mean ± SEM of the ERPs of women and men. The time period of significant topographic differences between women and men revealed by the TANOVA analysis (from 119 to 176 ms) is highlighted by the black line. The time periods of appearance of the microstate maps 2 (green line) and 3 (red line) -which respectively differed in terms of their onset and offset between women and men - are also shown for women and men.
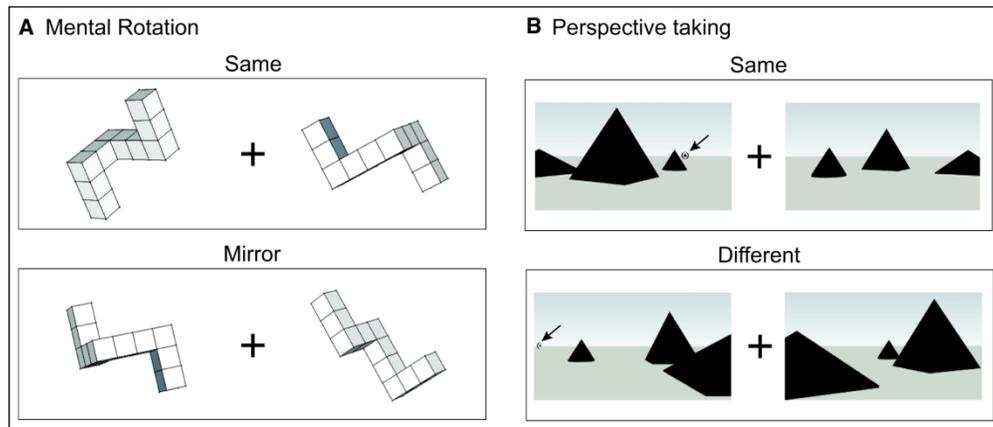
Figure 1: (A) In the MR condition participants had to judge whether the images on the right were the same (top) or different (bottom), (B) while in the PT condition they had to judge whether the images on the right coincided (top) or not (bottom) with the location of an observer, symbolized by an eye (arrow).
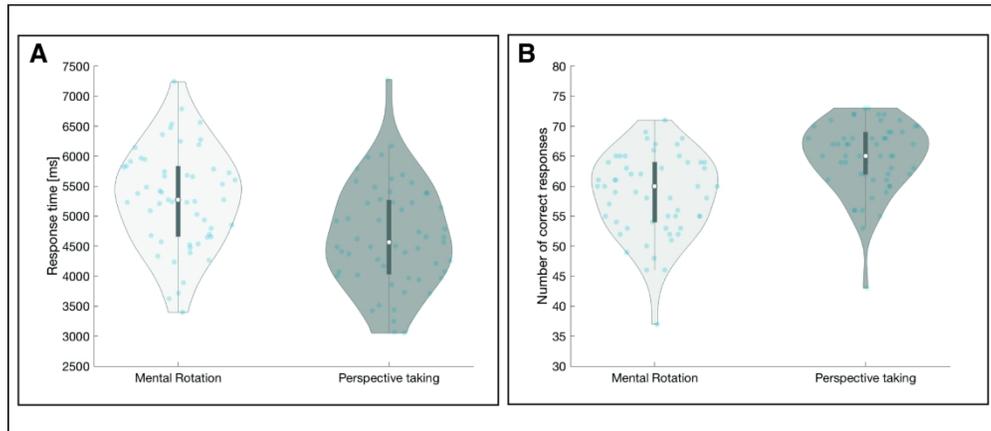
287x123mm (300 x 300 DPI)

Figure 2: Violin plot representation of the (A) response time and the (B) number of correct responses distributions for the MR and PT tasks. The light blue dots represent the original data and the violin plot outlines represent kernel probability densities of the original data. The white dot corresponds to the median value of the distribution, while the dark grey line illustrates the inter-quartile range between the 25th and the 75th percentile.
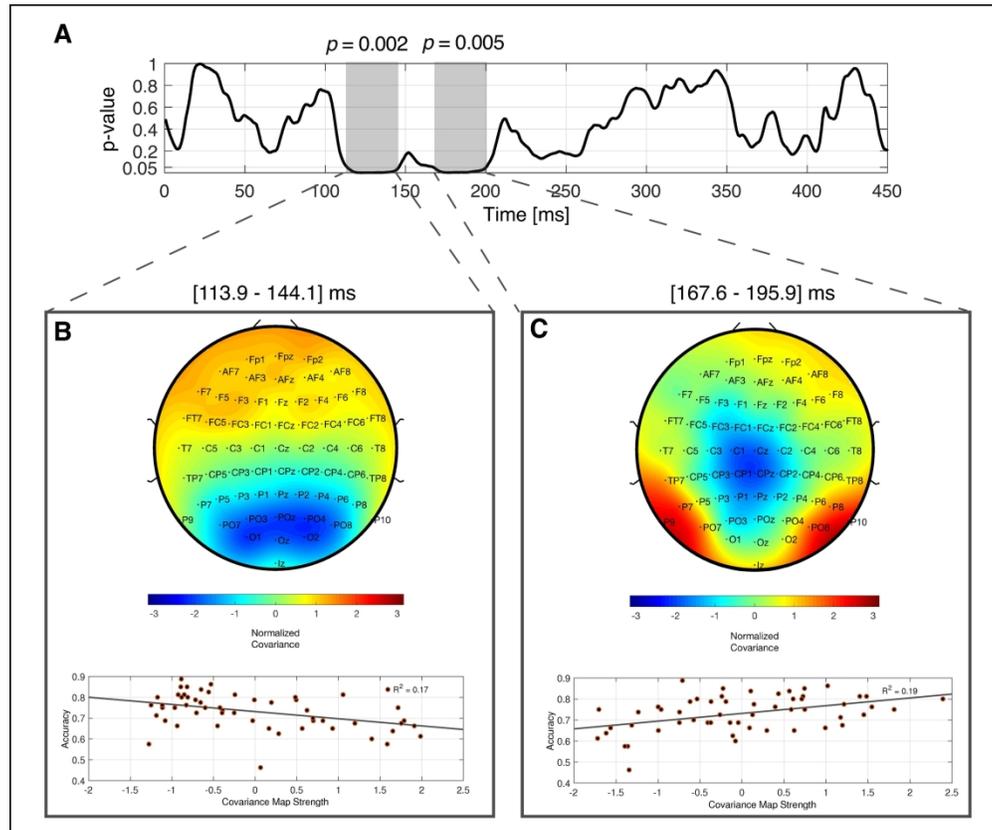
240x104mm (300 x 300 DPI)

Figure 3: Results of the TANCOVA analysis performed using the ERPs and the accuracy scores of the MR condition. (A) The p-value of the TANCOVA is plotted as a function of time. The grey shadowed areas highlight the time periods of significant covariation of ERP topographies and accuracy scores. (B and C) The mean topographies of the covariance maps observed from 113.9 to 144.1 ms and from 167 to 195.9 ms (top) are illustrated. The graphs below indicate the results of the linear regression analysis performed between their observed strengths in the ERPs and the accuracy scores.
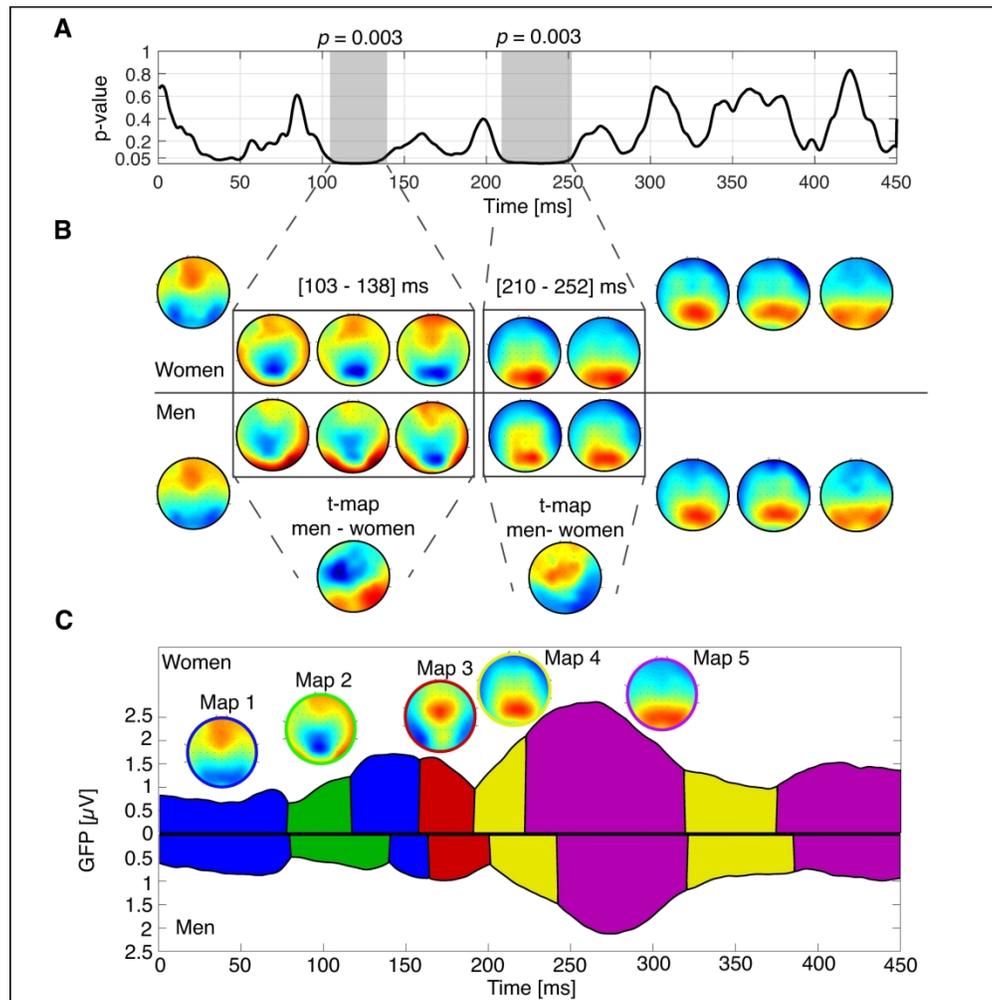
240x200mm (300 x 300 DPI)

Figure 4: TANOVA and microstate analysis results outlining periods of topographic differences of the grand-mean ERPs of the PT condition between women and men. (A) The p-value of the TANOVA analysis is plotted as a function of time. The grey shadowed areas highlight the time periods of significant differences between topographies of women and men. (B) Time course evolution of the ERP topographies of women (top) and men (bottom). Topographies within the time period of significant topographic difference are illustrated in gray boxes and displayed at higher time resolution (topographies' color bar is scaled between -3 and +3 □V); also shown are standard t-maps contrasting average topographies of men and women in the same time windows (t-map' color bar is scaled between -5 and +5). Positive (red) and negative (blue) t-values indicate more positive and more negative potential for men compared to women, respectively. (C) Topographies are shown for the five microstate maps obtained from the cross-validation procedure, along with their periods of appearance (coloured areas) displayed under the GFP (vertical axis) in the grand mean data of women (top) and men (bottom) as a function of time (horizontal axis).
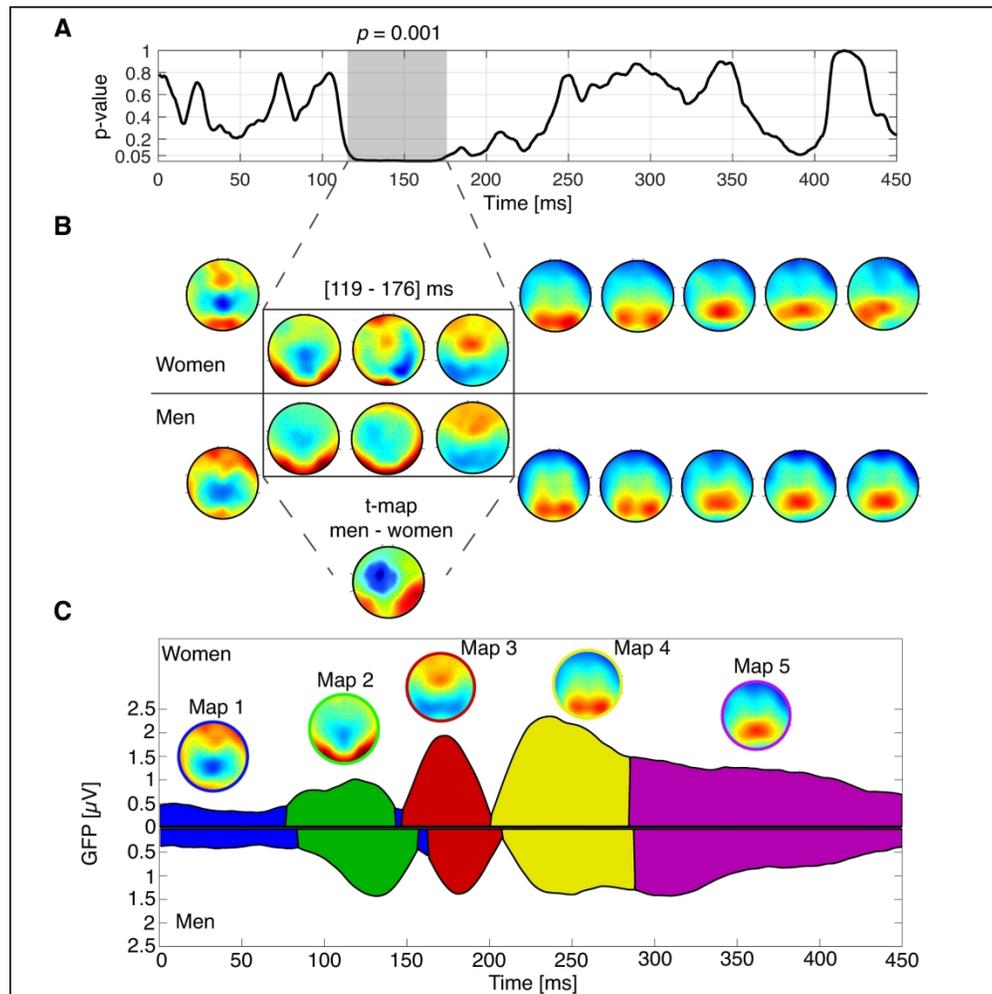
240x241mm (300 x 300 DPI)

Figure 5: TANOVA and microstate analysis results outlining periods of topographic differences of the grand-mean ERPs of the MR condition between women and men. (A) The p-value of the TANOVA analysis is plotted as a function of time. The grey shadowed area highlights the time period of significant differences between topographies of women and men. (B) Time course evolution of the ERP topographies of women (top) and men (bottom). Topographies within the time period of significant topographic difference are shown in the gray box and displayed at higher time resolution (topographies' color bar is scaled between -3 and +3 □V); also shown is the standard t-map contrasting average topographies of men and women in the same time window (t-map' color bar is scaled between -5 and +5). Positive (red) and negative (blue) t-values indicate more positive and more negative potential for men compared to women, respectively. (C) Topographies of the five microstate maps obtained from the cross-validation procedure, along with their periods of appearance (coloured areas) displayed under the GFP (vertical axis) in the grand mean data of women (top) and men (bottom) as a function of time (horizontal axis).
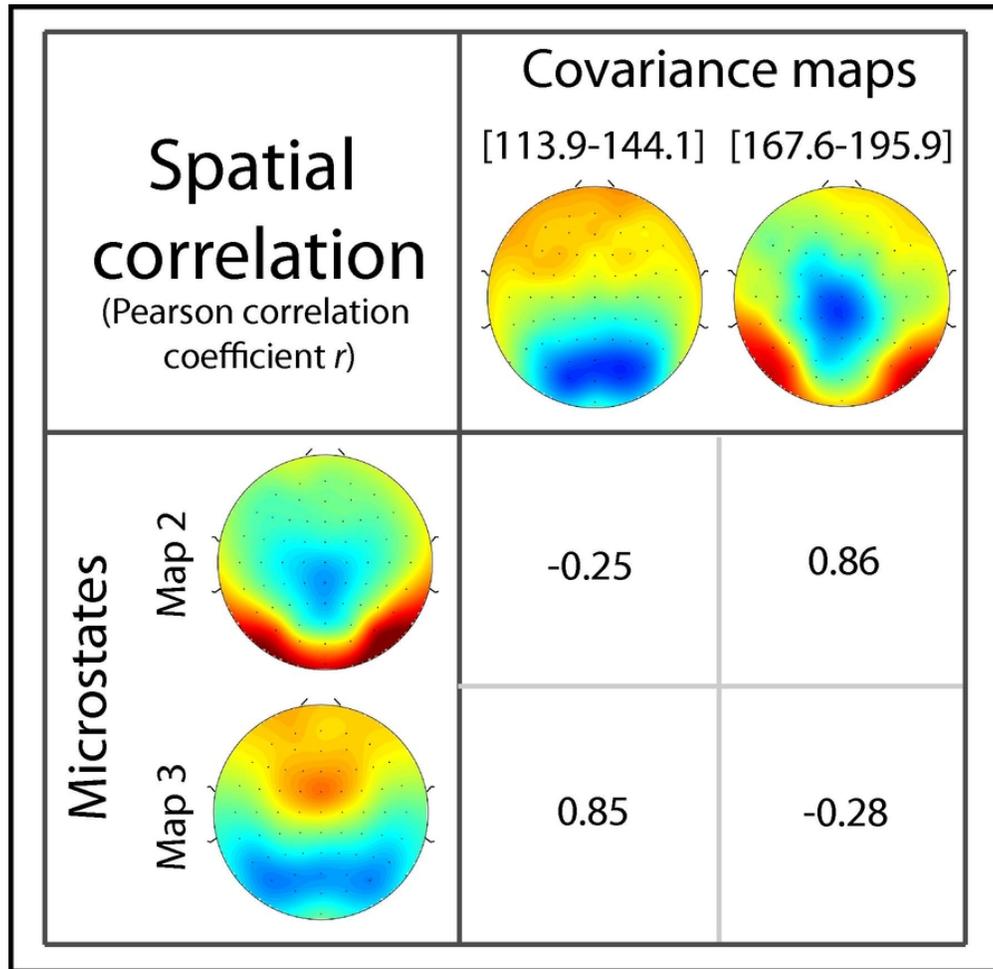
240x240mm (300 x 300 DPI)

Figure 6: Spatial correlation between the microstate maps 2 and 3 identified by the microstate analysis and the covariance maps revealed by the TANCOVA analysis in the MR condition.

75x73mm (300 x 300 DPI)